DOCUMENT RESUME

ED 066 726 24 CS 000 137

AUTHOR	Dansereau, Donald F.
TITLE	Analysis of the Reading Comprehension Process: The Development and Utilization of an Assessment Technique and the Preliminary Exploration of Individual Differences in Perceiving Syntactic Complexities. Final Report.
INSTITUTION	Texas Christian Univ., Fort Worth.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO	BR-1-F-055
PUB DATE	Jun 72
CONTRACT	OEC-6-71-0528(509)
NOTE	51p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Decoding (Reading); Reading; Reading Ability; *Reading Comprehension; *Reading Research; *Reading Skills; Reading Tests; *Sentence Structure; Word Recognition

ABSTRACT

The purposes of this research were to develop a new methodology for measuring comprehension and analyze individual differences in perceiving sentence difficulty. Written descriptions of commonly occurring visual scenes served as the stimuli in the new methodology. The reader's success in drawing a scene after viewing a written description of it served as a measure of his comprehension. Material was presented sequentially in a visual spatial mode or as a verbal description. The locations of scene objects were presented either in a random or systematic order, spatially proximal stimuli being presented contiguously in the systematic ordering. The results indicated that spatial presentation is better than verbal, and that systematic is substantially better than random. In addition, moderate but consistent relationships between task performance and reading ability were found. Two studies were conducted to explore individual differences in perceiving sentence complexity. Multidimensional scaling and other more traditional analyses showed a marked difference between good and poor readers in their perception of a variety of sentence characteristics. (Author/WR)

ED 066726

1-F-055
PA 24
NCERD
CS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Final Report

Project No. 1F055
Contract No. OEC-6-71-0528-(509)

Donald F. Dansereau
Texas Christian University
Fort Worth, Texas 76129

ANALYSIS OF THE READING COMPREHENSION PROCESS: THE DEVELOPMENT AND UTILIZATION OF AN ASSESSMENT TECHNIQUE AND THE PRELIMINARY EXPLORATION OF INDIVIDUAL DIFFERENCES IN PERCEIVING SYNTACTIC COMPLEXITIES

June, 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

(Regional Research Program)

FILMED FROM BEST AVAILABLE COPY

CS 000 187

ABSTRACT

The research can be divided into two categories: (a) development of a new methodology for measuring comprehension, and (b) analysis of individual differences in perceiving sentence difficulty.

Written descriptions of commonly occuring visual scenes serve as the stimuli in the new methodology. The reader's success in drawing a scene after viewing a written description of it serves as an objective measure of his comprehension. In an experiment with the scenes, we have presented the material sequentially in a visual-spatial mode or as a verbal description. The locations of scene objects were presented either in a random or systematic order, spatially proximal stimuli being presented contiguously in the systematic ordering. The results indicate that spatial presentation is better than verbal, and that systematic is substantially better than random. In addition, we have found moderate, but consistent relationships between task performance and reading ability. Plans for extending this methodology are discussed.

In the second category of research, two studies were run to explore individual differences in perceiveng sentence complexity. Multidimensional scaling and other more traditional analyses showed a marked difference between "good" and "poor" readers in their perception of a variety of sentence characteristics (for example, syntactic complexity, familiarity, and meaningfulness).

Final Report

Project No. 1F055

Contract No. OEC-6-71-0528-(509)

Analysis of the Reading Comprehension Process: The development and Utilization of an Assessment Technique and the Preliminary Exploration of Individual Differences in Perceiving Syntactic Complexities

Donald F. Dansereau

Texas Christian University

Fort Worth, Texas 76129

June, 1972

The research reported herein was performed persuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

Table of Contents

		Page Numbers
List of Tables		2
List of Figures		3
I.	Overview	4
II.	Research with the "Scene" Methodology	4
	A. Introduction	4
	B. Method	12
	C. Results	14
	D. Discussion	15
	E. Conclusion and Planned Extensions	16
III.	Research Assessing Individual Differences in Perceiving Sentence Difficulty	21
	A. Introduction	21
	B. Experiment 1: Perception of Syntactic Structure	22
	C. Experiment 2: The Effect of Reading Ability and Exposure Time on Judgment of Complexity, Familiarity, and Comprehensibility	33
IV.	References	44

LIST OF TABLES

Table number	Title	Page number
1	ANOVA Summaries	36
2	Mean Comprehensibility Judgments for Variations in Meaningfulness	39
3	Correlations of Comprehensibility Judgments with Sentence Difficulty	40
4	Percentages of Significant Consistency Correlations	41
5	Judgment Intercorrelations	43

LIST OF FIGURES

Figure number	Title	Page number
1	Street Scene	10
2	Pier Scene	11
3	Hierarchical Scene	19
4	CLIFTON-ODOM Transformation Results	25
5	Weighting of the Passive Dimension vs. Weighting of the Question Dimension	26
6	Weighting of the Paraphrase Dimension vs. Weighting of the Question Dimension	27
7	Weighting of the Negative Dimension vs. Weighting of the Question Dimension	28
8	Weighting of the Paraphrase Dimension vs. Weighting of the Negative Dimension	29
9	Weighting of the Passive Dimension vs. Weighting of the Negative Dimension	30
10	Weighting of the Paraphrase Dimension vs. Weighting of the Passive Dimension	31
11	Mean Comprehensibility Judgments as a Function of Sentence Difficulty for Time 1	37
12	Mean Comprehensibility Judgments as a Function of Sentence Difficulty for Time 2	38

I. Overview

The activities accomplished under OE Grant No. OEC-6-71-0528-(509) can be divided into two general categories: (a) assessment and utilization of a new methodology for measuring comprehension and (b) empirical analysis of individual differences in perceiving sentence difficulty. In addition to these activities, we also have continued to collect and annotate articles related to readability, reading comprehension, and psycholinguistics. To date, 250 references have been accumulated, 150 of these have been annotated, and 150 have been reproduced for future usage.

The first category of work, which was the primary aim of the initial proposal, received the bulk of our efforts. Analysis of a second assessment methodology, utilizing multidimensional scaling, was proposed, but work on this approach was discontinued due to unforeseen difficulties in its implementation (see the January 15, 1972 Progress Report for details). We have just recently discovered ways to circumvent the problems with this latter technique and have included it in a recent proposal to extend our present work (see Section I E 2 of this document).

The second category dealing with perceived sentence difficulty, was initiated to explore an area which as a result of our review of the literature, we felt was critical to understanding the comprehension process. Also it appeared that multidimensional scaling would have much stronger potential in this area.

We have three major studies to discuss in this report. A number of smaller exploratory studies have been completed but their import is limited. Their main purpose was to provide information relevant to optimizing the designs of the major studies so they will not be explicitly dealt with in this report. For clarity, the three studies will be subdivided into the two major categories.

II. Research with the "Scene" Methodology

A. Introduction

Bormuth (1970) in extrapolating from his work on middleclass suburban schools has estimated that 51% of the students graduating from high school are illiterate with respect to the average materials used in their instruction and 46% are illiterate with respect to the average passage from the news media. Even though the assessment techniques used by Bormuth may be questionable (see the discussion of the cloze technique in a subsequent paragraph), his estimates would be a basis of concern even if they were substantially discounted.

In discussing his findings on literacy, Bormuth (1970) has pointed out the need for a model of comprehension. "In order to develop a reasonably effective theory of (comprehension) instruction, it is necessary to have fairly accurate descriptions of how these (comprehension) processes link together to produce a complex behavior. Clearly, any effort to build such a theory would have to draw upon linguistic theory of the structure of language and upon psychological knowledge of how to analyze complex mental processes." Delineation of the underlying comprehension processes and their interactions will require the development of adequate techniques for the manipulation of variables hypothesized to influence comprehension and the subsequent measurement of the magnitude of their effects.

As will be seen in subsequent sections, the methodologies developed to explore the comprehension processes have not been adequate. This research represents a step toward developing and examining a methodology which we believe will eventually remedy this situation.

(1) Presently used Methodologies for Assessing Comprehension

To set the stage for description of this methodology, we will briefly discuss the shortcomings of presently used techniques in the context of a particular conceptualization of the comprehension process.

(a) A description of the comprehension process and its relationship to retention.

It is necessary to distinguish comprehension from transcription (typing and reading aloud) and rote memorization, in both of these cases there is a one to one relationship between the input and the subsequent output. Although the ability to transcribe is necessary for successful comprehension, the critical process in comprehension involves the translation of the stimulus material into an appropriate internal representation; abstraction of theme or gist falls into this category. It is proposed that the essence of comprehension is the abstraction of relationships. These may be relationships between words or combinations of words in a sentence, words or combinations occurring in different sentences, as well as higher order relationships between relationships derived at lower levels. For example, a relationship between Ann, dog and Tom may be abstracted from the sentence: "Tom's dog bit Ann" and this relationship may, in turn, be related to the one derived from a later sentence: "Ann hit Tom with her wounded hand." The nature of these relationships cannot be specified at present, but it appears that they are multidimensional and may be adequately represented in a spatial memory model. Stored

hierarchies of such relationships (that is, compound relationships between progressively lower order relationships), along with those relationships linking presently read material to that which was previously stored (that is, the integration of new knowledge with old through the development of new relations), would represent the comprehension of a passage. It should be noted that this description of comprehension differs only slightly from that of Frase (1969a, 1969b) and is presented as a framework in which to view previous and future research and not as a serious model.

Although many researchers have called for methodologies which analyze comprehension independently of memory processes (for example, Carroll, 1971), it is clear that storage of the abstracted relationships both during and after the comprehension process is critical to successful performance in an educational setting. In fact, some sets of relationships between concepts presented in a passage may be easily abstracted, but due to any number of variables (for example, lack of concreteness, irrelevancy, disorganization) may be only subsequently retained for a minimum length of time. In addition, since during comprehension some effort to store the abstracted material presumably takes place, there is a type of time sharing which may be reflected in retention difficulties. For example, a complex passage (due to unfamiliarity, syntactic and organizational complexity, etc.) may require substantial processing (both syntactic and semantic) in order to provide accurate comprehension. This increased effort may reduce the amount of processing devoted to storing the abstracted relationships. With simpler passages a more equitable balance between understanding and storage processes may be maintained. Therefore, differences between complex and simple passages may not appear under immediate recall conditions, but may be reflected only in retention difficulties over time. Naturally, an appropriate methodology will allow for the analysis of both comprehension and retention processes as well as their interactions.

(b) Shortcomings of presently used methodologies for exploring comprehension and retention.

Although there have been a variety of methods for measuring comprehension and diagnosing processing difficulties, there have been only a few classes of techniques which have been applied to multiple sentence connected discourse.

(i) Free recall. This technique, which requires the S to reproduce the essence of the presented material in his own words, has probably the greatest face validity of any of the methods. The major difficulty is one of scoring the Ss output in an efficient, economical fashion. To date, this

problem still remains unsolved. Crothers (1971) is developing a procedure for measuring paragraph structures which should eventually be applicable to comparing paragraph input with Ss' outputs. Until this approach is perfected, the free recall technique remains very difficult to use in both testing and experimental situations.

(ii) Multiple choice tests. The most common technique used in previous research on comprehension, and in practical testing situations, has been the multiple choice test. This approach has been used in both comprehension and readability studies. Comprehension has been studied by factor analyzing the responses to questions which purportedly measure different components of comprehension ability. For example, a typical set of questions to be administered following the reading of a passage, would be designed to test: word knowledge, ability to manipulate ideas and concepts in relation to one another, ability to grasp the author's expressed ideas, ability to identify the writer's intent or purpose, ability to follow the organization of a passage, and knowledge of literary devices and techniques (Davis, 1944).

The results of a number of analyses of this type (Davis, 1944; Conant, 1942; Thurstone, 1946; Hall and Robinson, 1945; Stoker and Kropp, 1960; Davis, 1968; Singer, 1969) have been equivocal. It is unclear from this research whether there is a single general factor underlying comprehension or multiple factors. One major difficulty with this research, which may to some degree account for the ambiguous results, is the apparent lack of valid and replicable comprehension test questions. In effect, different researchers do not necessarily agree on what factors specific questions do measure, nor do they necessarily agree on the level of difficulty exhibited by each question (see Lorge, 1939; Bormuth, 1966; and Bormuth, 1970). This latter problem, concerning the difficulty of the test questions, also casts doubt upon the validity of the readability findings.

(iii) The cloze technique. The cloze technique developed by Taylor in 1953 has been used extensively in analyzing the readability of passages. This method generally involves the replacement of every fifth word in a passage with an underlined blank. Subjects, who have not previously read the passage, are asked to write in the words they think were deleted and their responses are scored correct when they exactly match the deleted words. A number of studies (Taylor, 1953; Bormuth, 1967; and Rankin, 1965) have shown that the cloze technique is a valid and reliable measure of comprehension difficulty when responses to post-reading questions serve as

criteria. For example, Bormuth (1962) found a .77 to .87 correlation range at the fourth-grade level when individual cloze and multiple choice comprehension tests scores were correlated over the same material.

This technique has been used in lieu of post-reading questioning to determine the effects on reading difficulty of quantitative and qualitative differences in linguistic properties at the level of: words, clauses, sentences, and longer passage segments (Bormuth, 1966; Ruddell, 1965; MacGinitie and Tretrak, 1969; Bormuth, et al, 1970). A large number of these linguistic properties (for example, word length, sentence length, and number of clauses within a sentence) have high and predictable correlations with reading difficulty as measured by the cloze technique. Hopefully, knowledge of the linguistic factors which facilitate or inhibit comprehension will allow us to construct graded reading material, diagnose reading difficulties, and infer processes underlying comprehension. The cloze technique, although objective, may be misleading in that it is difficult to determine exactly what is being measured. Validation of this technique, as discussed previously, has been accomplished by comparing its results with that of multiple choice test questions. If the validity of the test question approach is in doubt, as users of the cloze technique have themselves pointed out, then the validity of the cloze technique is also in doubt.

Even if it is assumed that this technique is a valid indication of comprehension difficulty at some level, its usefulness would appear to be limited to the abstraction of relationships which are in close proximity within the passage. In fact, MacGinitie (1961) found that cloze items are statistically independent when surrounded by five words of context; that is, uninterrupted context five words on either side of the target did not help in the restoration of the missing word. This would imply that the technique is measuring difficulty over small segments of text, and would be insensitive to the abstraction of higher order relationships, such as those concerned with paragraph organization. In addition, the typical usage of the cloze technique would preclude the analysis of the retention aspects of comprehension.

(2) Description of the Scene Methodology

As an initial step toward producing a more efficient methodology for analyzing and testing comprehension and retention processes, common, aerial-view scenes were developed (Figures 1 and 2). One purpose of using physically realizably concepts was to provide an objective, alternate format in which a reader can express his comprehension of a passage. For example, a reader

of a set of instructions for constructing a crib, can show his comprehension of these instructions by answering questions about the material or by actually attempting to build the object. Although motor skills may confound the results, this latter procedure should provide an objective test of the individual's comprehension of the passage or set of passages. In a like fashion, descriptions of movement, objects, visual scenes, etc., should contribute reasonable alternative measures of comprehension.

A second advantage to using physically realizable concepts, especially scenes, is that the basic information (deep structure) can be presented in alternative ways - either as a verbal description or as a picture. These alternative modes of presentation allow the experimenter to separate language decoding aspects from other comprehension processes.

Previous work with physically realizable concepts has been very limited. A number of studies have used the ability of Ss to follow directions as a measure of comprehension (Jones, 1966; Brown, 1955; Shipley, Smith, and Gleitman, 1967). These studies have been limited to using only a few sentences and are hampered by the lack of an alternative mode of presentation.

Other studies have tested comprehension by requiring verification against pictured referents (Gough, 1965, 1966; Slobin, 1966). This procedure is most closely aligned with the "scene" methodology, with three exceptions: (a) picture verification has been applied only with single sentences, (b) the verification process allows for guessing which impedes the validity of any single response, and (c) presentation in alternative modes has not been explored (verbal vs. visual). The "scene" methodology represents an extension of these studies on all three dimensions.

To date, five scenes have been constructed, these consist of aerial views of an airport, pier, street intersection, farm, and grocery store plus a modified living room scene for practice purposes (examples are shown in Figures 1 and 2). These scenes, which consist of seven objects on a three item background, have been equated in a variety of ways, thus permitting generalization through replication. In addition to the scenes themselves, we have developed standard verbal descriptions of the scene objects and their locations. The number of words, the number of sentences, and the sentence structure have been equated over all descriptions

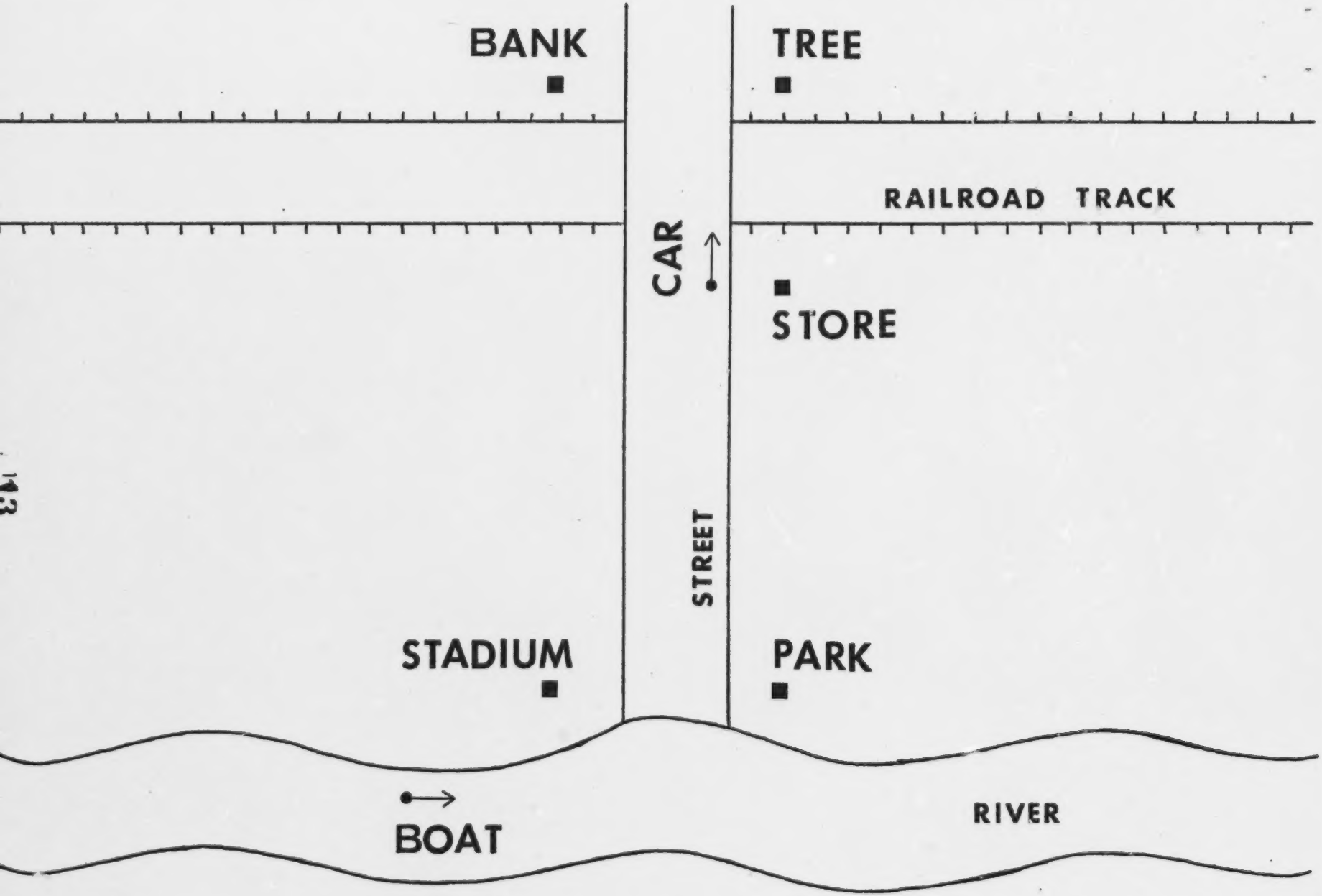

BANK
TREE
RAILROAD TRACK
CAR
STORE
STREET
STADIUM
PARK
BOAT
RIVER

FIGURE 1

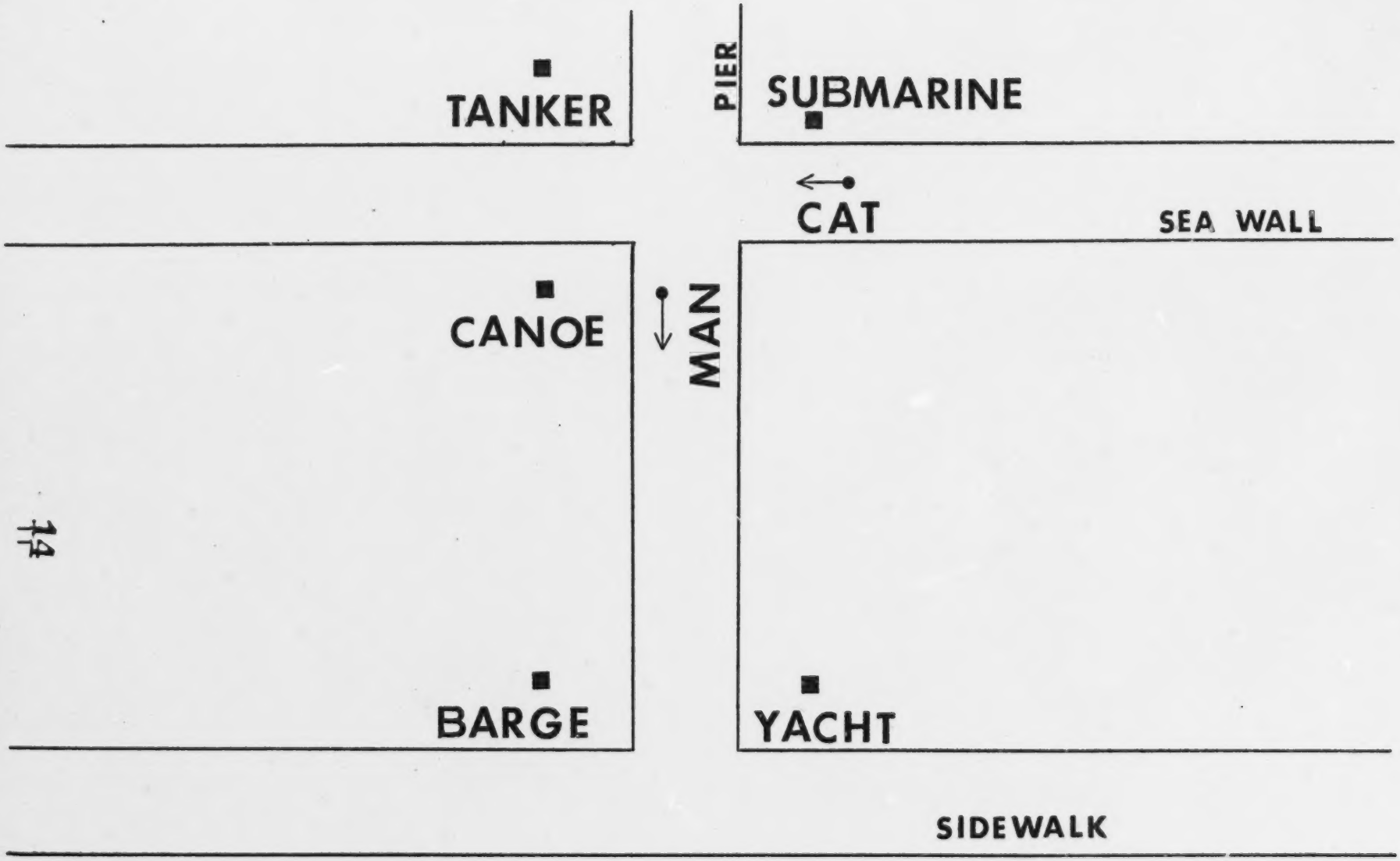

TANKER
PIER
SUBMARINE
CAT
SEA WALL
CANOE
MAN
BARGE
YACHT
SIDEWALK

FIGURE 2

(3) Preliminary Assessment of the "Scene" Methodology

The philosophy underlying the present research is that a methodology can be most efficiently tested, even at the initial stages, by using it in an experiment designed to determine the empirical relationships between meaningful variables. With such an approach, there are, in addition to the resulting methodological developments, potentially strong empirical consequences which may lead to new implications for a theory of comprehension.

The specific independent variables investigated in this research and their operational definitions are as follows:

(1) Syntactic complexity: This is operationally defined as the voice of the sentence or passage. Greenough and Semmel (1969) suggested, as have others, that active sentences because they are most frequently encountered are easier to comprehend and recall than passive sentences.

(2) Order of presentation: The locations of objects in a visual scene can be described either in an order based on a logical progression through the scene using spatial proximity as the criterion, or in a random order. It was reasoned that the logical, systemmatic order should facilitate comprehension and recall.

(3) Mode of presentation: The locations of the objects can be conveyed using either written descriptions or through the use of diagrams. If reading difficulties have to do with the ability to translate information from a written code into a visual code (in the case of descriptions of visual scenes) then good and poor readers should do equally well on the diagrammatically presented information, but not with the written descriptions.

(4) Reading ability: This variable was defined in terms of the total reading score on the Cooperative English Reading Test which is administered to all freshmen entering Texas Christian University.

B. Method

Subjects

Eighty students, fulfilling an Introductory Psychology course requirement at Texas Christian University served as Ss for a total of one hour per individual. The forty Ss with the lowest "total reading" percentile on the Cooperative English Reading Test were arbitrarily assigned to the poor reading group. The remaining forty Ss were designated as members of the good reading group. The range of scores in the poor reading

group was from the 7th percentile to the 68th percentile and the range for the good readers was from the 75th to the 99th percentile. Good and poor readers were randomly assigned to four experimental treatment groups such that each group was composed of ten good and ten poor readers. In addition, Ss were assigned to groups with the constraint that the mean total reading percentiles for the four groups were approximately equal.

Stimulus Material

Aerial views of six familiar scenes (a city intersection, a farm, a pier, an airport, a living room, and a grocery store) presented either in the diagrammatic format illustrated in Figures 1 and 2 or in the form of written descriptions comprise the stimulus set from which the stimuli used in the various phases of the experiment were selected. Each scene consists of a background which is standard for all of the scenes except for labeling, plus seven objects superimposed on that background. For each scene, three components of the background are labeled.

Procedure

Each S, regardless of experimental treatment group assignment, was presented with four scenes (one of which was for practice) via black and white slides in the following manner. First the background with three labeled components was presented (either in diagrammatic or written format) for 24 seconds. For the written format this corresponds to a reading rate of 150 words per minute. Next, each of the seven scene objects was presented for eight seconds in conjunction with the background. One minute after presentation of the last object the Ss were asked to list the seven objects and the three labeled background components in the order in which they were presented. Upon completion of this task they were asked to locate and label the seven objects and label the three background components on an unlabeled background. The dependent measures corresponding to these two tasks are total items correctly recalled, and total items correctly located and labeled.

The four experimental groups, each containing good and poor readers represent the eight possible cells of a fully crossed 2x2x2 design with reading ability, mode of presentation (diagrammatic or written), and order of presentation (systematic or random) as factors. In order to investigate the effect of syntactic complexity upon retention, the written groups, both systematic and unsystematic were presented with descriptions of the scenes such that every other sentence (each sentence corresponding to the location of one object on the background) was in the active voice and the remaining sentences were in the passive voice. Thus for the written group only, an additional within-subjects factor was introduced.

In addition to this first task which will be termed "Phase I" of the experiment, a second task termed "Phase II" was required of the Ss in the written group. These Ss, upon completion of "Phase I" were given a set of 3x5 cards each containing one sentence. The set of cards, taken together, described two additional scenes, with order of presentation again being either systematic or random. For each S, one of the two scenes was presented entirely in the active voice and the second entirely in the passive voice. Ss were provided with unlabeled backgrounds and told to read the cards and after reading each card to locate and label (by initial letter only) the appropriate scene object on the blank background as quickly as possible. Total time to locate and label the elements of each scene was the dependent measure.

The second phase of this experiment can be conceptualized as a 2x2x2 fully crossed design with syntactic complexity (active vs. passive voice) as a within S factor. Syntactic complexity and the two scenes were appropriately counterbalanced to insure that the two would not be confounded.

C. Results

Phase I

2x2x2 fixed effects analyses of variance done on the two dependent measures with reading level, mode of presentation, and order of presentation as factors indicated the following.

(1) The diagrammatic presentation mode resulted in significantly better recall than did the written mode ($p < .01$).
(2) The systematic presentation order resulted in significantly better recall than did the random order ($p < .05$).
(3) Mean items recalled and mean items located and labeled were consistently greater for good than poor readers in comparable treatment conditions but even so, this factor as well as all interactions were non-significant.

2x2x2 analyses of variance done on the two dependent measures for the written group, with presentation order, reading level, and semantic complexity as factors, indicated that only the presentation order variable was significant, with systematic leading to better recall than random ($p < .05$). Again, good poor differences, though not significant, were in the expected direction.

A correlation of .50 was found between each of the two dependent measures in Phase I and the total reading percentile on the Cooperative English Reading Test. This suggests a weak but significant relationship between these measures.

Phase II

A similar 2x2x2 analysis of variance done on the same three factors in Phase II revealed no significant effects at all. Again, good poor-differences were consistent and in the predicted direction.

D. Discussion

The lack of any significant effect involving the reading level variable (good vs. poor) was surprising. The reasons that the consistent differences between good and poor readers were not statistically significant may be due to: (1) lack of sensitivity of the dependent measures and/or (2) lack of consequential differences in reading ability between the good and poor readers. It was felt that correlational analysis would provide more detailed information on the relation between reading scores and retention of the scene information. The small but significant correlation of 0.50 suggests that reading test scores and performance on the experimental task are to some extent measuring the same underlying ability. Only further work with this methodology suitably extended so that the dependent measures are appropriately sensitive will properly reflect the magnitude of this relationship.

The insignificance of the syntactic complexity variable could also be due to the insensitivity of the dependent measures, since as we have already stated, differences between the retention of active and passive sentences have been demonstrated in previous studies. It was hoped that syntactic complexity would interact with reading test scores, showing a greater disparity in performance between good and poor scorers with the more complex passive sentences than with the active. Unfortunately, it appears that the dependent measures were not sufficiently sensitive to record such an interaction. Further work with more refined measures will aid in determining the effects of this variable. The second category of studies accomplished during the grant period bears on this issue (see Section III).

The significant order effect was consistent with the findings of Oakan, Wiener, and Cromer (1971), however, the lack of any interactions of this variable with reading level is inconsistent with their hypothesis that poor readers, because of organizational disability are less able to comprehend information that is not already well organized. It should be noted, however, the insensitivity of the dependent measures in this research may well be masking this interaction as well as other expected effects.

The mode of presentation effect was surprising in that it occurred with both dependent measures - recall of scene items and locating and labeling. While it is obvious that a visual presentation is a more direct method of conveying the spatial information, it was not anticipated that it would so greatly facilitate the simple recall of the items names. In retrospect, this can be explained in one of two ways. First, if the visual presentation facilitates a non-verbal storage of the object locations, then more verbal storage would be free to store the item names. Second, if the visual presentation conveys the spatial aspects of the scenes more quickly, then the Ss would have more time in which to rehearse the item names.

E. Conclusions and Planned Extensions

At this point, we are forced to conclude that the utility of the "scene" methodology as a reading assessment technique is somewhat limited. It appears that the task, as presently structured, is not powerful enough to detect subtle, theoretically meaningful differences in performance. We feel that we have found a solution to this lack of sensitivity and will discuss it in sub-section (2). Prior to this discussion, however, we would like to describe some plans for future work in an area in which we feel the presently developed scenes will be extremely useful.

(1) Temporal Organization (Order of Presentation)

In many practical cases, there is a pressing need to present information in a way which will maximize performance in tasks requiring comprehension, retention, or utilization. In pursuit of this objective, much research has focused on the effects of order of presentation (or temporal organization). Although there has been a large amount of research on this variable, especially in the domain of simple retention studies, the results have not been integrated sufficiently nor have the research methodologies been powerful enough to permit substantial generalization and application.

Our work in pattern recognition, human memory, and reading comprehension has required us to look very carefully at the effects of stimulus ordering in these tasks. As a result of this research and of a preliminary review of the literature on stimulus ordering in a broad range of tasks, we have delimited an area of study which we feel now deserves substantial attention. The area lies between simple retention studies (serial list learning, free recall, and recognition memory) and the more complex tasks of meaningful prose comprehension and retention. Previous research on stimulus ordering has focused heavily on simple retention studies; a good deal

is known about the serial position effect, isolation effects, and conceptual clustering as they occur in these tasks. Furthermore, a number of sophisticated models - some quite successful - have been developed to account for these phenomena.

Once we move out of this area, however, to tasks involving more complex underlying stimulus organizations (for example, hierarchial structures) and more complex processing, sophistication of the work decreases. In particular, this decrease is quite marked in the transition from simple retention to prose comprehension and storage. It appears that two kinds of activity are needed in order to extend our knowledge of temporal organization to more complex tasks. First it is necessary to specify the underlying stimulus organizational structure which must be sufficiently transmitted in order to produce effective performance in these tasks. This organizational structure will generally consist of a set of items and their interrelationships associated with a particular content area (for example, biological taxonomy). Structures can take the form of lists, hierarchies, and networks plus a variety of higher order conbinations of these simple structures. The ultimate requirement of this phase is a taxonomy or organizational structures within which a wide variety of tasks and content areas can be categorized.

In the second kind of activity, each organizational structure must be studied separately. The effects of varying the temporal organization of the information would be assessed and the results integrated into a model of human processing.

Since the "scene" stimuli have a two-dimensional network structure which is more complex than list structures normally used in simple retention studies, they provide a good basis for initiating this second type of activity. As stated earlier, the results from our "scene" experiment indicate that spatial presentation is substantially better than verbal, and that systematic is far better than random regardless of the mode of presentation. These results raise a number of questions which we have proposed to explore (a proposal has been submitted to the Office of Naval Research).

(a) Does pythagorean proximity provide the most efficient means of temporally organizing the material? Alternative orderings which will be explored are based on: "city block" proximity and presentation of various features of the Gestalt figure, such as presenting corner items prior to middle items, et cetera.

(b) What aspects of the human processing system determine efficient ordering in this task? The results from question one above will be analyzed with this in mind.

(c) What are the effects of expanding the number of stimulus items and manipulating the background configuration? The generalizability of the previous results will be tested.

(d) How pervasive are the effects of individual differences? Other measures of individual difference will be included in the proposed studies.

Answers to these questions will presumably give us some indication of temporal organization effects in simple two dimensional networks.

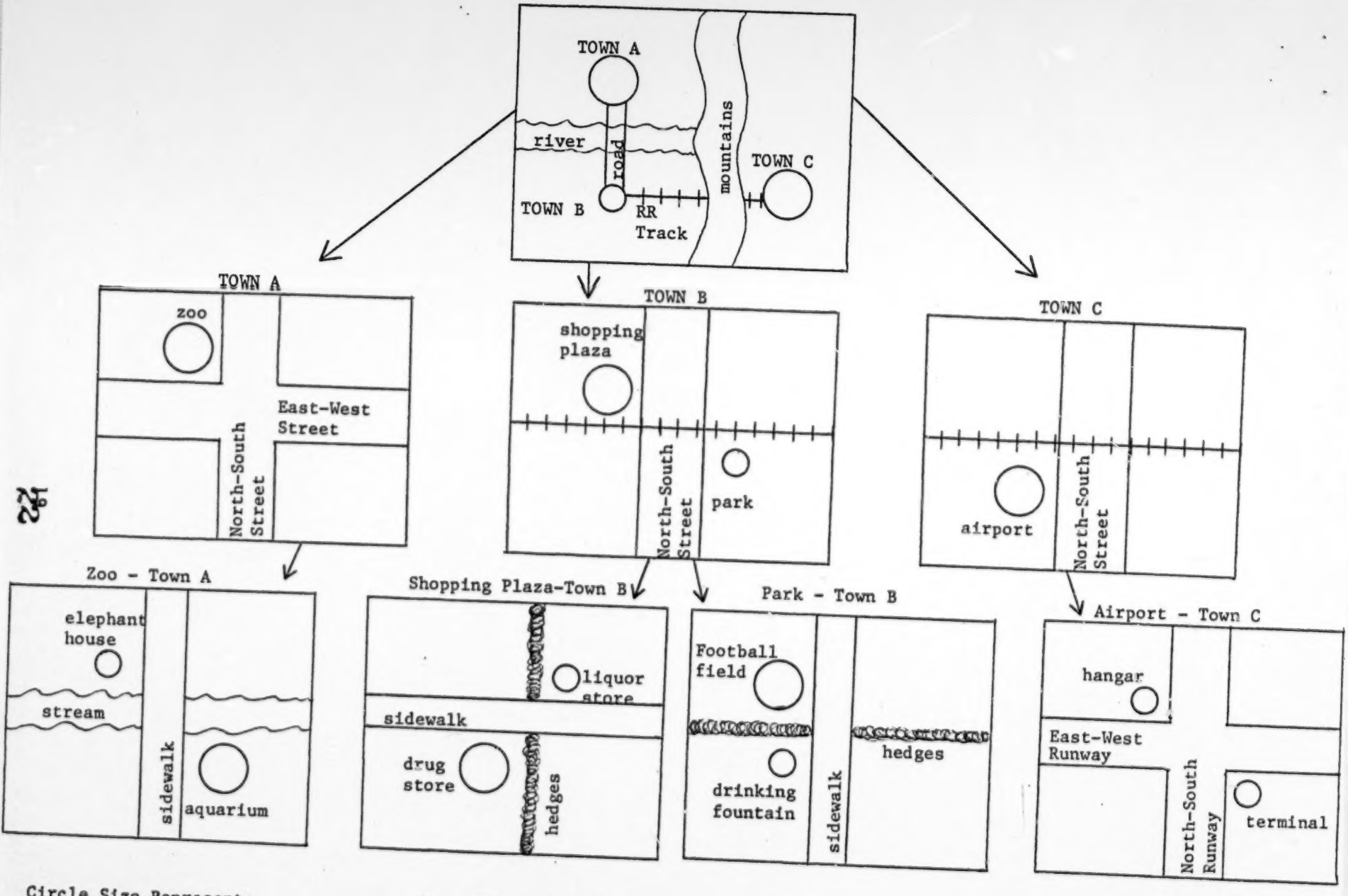
(2) Development of Hierarchical Scenes

Hierarchical structures predominate in educational settings; the structures of textbooks, courses, and curricula are almost always hierarchical. In addition, although most physically relizable concepts are not inherently hierarchical in nature, we often find it convenient to describe them hierarchically (for example, instructions for games and filling out tax returns). In the theoretical domain, investigators have hypothesized that cognitive plans for problem solving and action are hierarchically organized (Miller, Gallanter, and Pribram, 1960), and that semantic memory consists largely of interlocking hierarchies (Miller, 1969; Collins and Quillian, 1969, 1970; Kintsch, 1970; Mandler, 1967).

Given the apparent importance of such structures, it is surprising that only a few studies have been directed toward analyzing the comprehension and/or retention of material organized in this fashion. Bower, Clark, Lesgold, and Winzenz, (1969), Rourke, 1971, and Segal, 1969 used hierarchies of words in free recall experiments and showed that the hierarchical organization "built into" a set of words to be learned can be recovered in the recall organization. However, none of the studies we are familiar with have manipulated structure or presentation variables, nor have they extended the free recall findings to the area of prose comprehension and retention.

In order to fill this gap and to provide a more sensitive comprehension assessment methodology, we have preliminarily developed hierarchical "scenes" (for example, see Figure 3) and have proposed to further develop and utilize these stimuli in theoretically meaningful experiments (proposals have been sent to Child Health and Human Development, National Institute of Mental Health, and the National Science Foundation).

The stimuli, which will be presented both diagrammatically and in the form of verbal descriptions, will reflect various conditions of hierarchical organization in the information to be presented. The scenes will be constructed to the specifica-



Circle Size Represents
Relative Object Size
Within Each Diagram

Diagrammatic Representation of a
Hierarchical Organized Scene

This Series of Diagrams
will Require a 600-800
Word Description

FIGURE 3

tions of each individual experiment. They will be analogous to the one shown in Figure 3, but naturally will deviate somewhat depending on the particular experimental manipulations employed. In addition to these aspects of the methodology, several other properties may be noted.

The scenes will be artificial to minimize subject differences based on previous experience and to permit construction of alternate forms in which the stimuli are identical except for the object names. These alternate forms will be used to provide replications of each experimental condition.

The scenes also exhibit several other manipulable properties. In particular, there will exist a number of dimensions along which subjects may structure their memories; these dimensions include object size (represented by the size of circle indicating its location), spatial proximity, interconnectedness (via roads, et cetera), and object class (type of item, such as business, recreation, et cetera). These dimensions can be arranged so as to be correlated with the locations of the objects in the hierarchical structure or independent of the locations. Presumably, having a number of dimensions correlated with the hierarchical structure will facilitate the use of this structure in memory.

With regard to the verbal description of each scene, the sentences will be generally in one of two forms:

"'A', which is a large town, is located immediately to the west of the mountains and just north of the railroad tracks"; or

"Immediately to the west of the mountains and just north of the railroad tracks is the large town of 'A'."

Since we are not explicitly interested in sentence processing we will not attempt to manipulate variables of sentence structure, but will use the two forms intermixed randomly to provide variety throughout the passage.

The conversion of a scene into verbal description is straightforward, although the resulting prose is hardly elegant. We are not entirely satisfied with our present method of conversion into verbal form and plan to explore alternative methods, including techniques to obtain representative samples of text produced by persons asked to describe the scenes.

As with the previous, non-hierarchical scenes, the S will locate and label the objects and the components of the background. An unlabeled background will be provided for this purpose. A variety of scoring procedures will be used to assess performance on this task; these include pythagorean

distance between objects in the original and in the response "scenes", and absolute number of correct items. This free recall of information should provide a direct measure of how well Ss comprehended and retained the original material.

Multidimensional scaling will be used to further assess the organization of the presented material in memory. With this technique (see Torgerson, 1958, for a more complete description), Ss will be asked to judge the relatedness of all possible pairs of scene objects. The resulting data will be analyzed by the INSCAL program (Carroll and Chang, 1970). The output from this program will be an inferred multidimensional representation of the Ss' memory structures. Previous work with this technique (Dansereau, Fenker, and Evans, 1970), has shown it to be of substantial value in determining how material ss stored. In particular, important dimensions of storage can be delimited.

In order to provide a basis for concurrent validation of the assessment of the inferred memory structure, the S will be asked to access the stored material along a number of dimensions. Questions such as, "name the three largest objects in the scene" or "name the three most proximal objects" will be asked. Latencies and errors in responding can then be used to draw inferences as to the nature of the internal structure and to detect changes in this structure as the result of experimental manipulations. Findings with this approach should be strongly related to the results of a multidimensional scaling analysis.

III. Research Assessing Individual Differences in Perceiving Sentence Difficulty

A. Introduction

One aspect of the comprehension process which we feel deserves attention is that of on-going decision-making. It is hypothesized that a good reader can vary his processing mode to suit the difficulty and importance or relevance of the material. The basic decisions are concerned with determining the appropriate time to shift from one processing mode to another. Decision to scan, concentrate, assimilate, et cetera, if made appropriately, produce greater reading efficiency. These decision processes have been studied only in superficial or applied experimental contexts under the label of "reading flexibility" (Bowman, 1966; Braam, 1963; Braam & Berger, 1968; Laycock, 1955; McDonald, 1965). In general, very little attention has been paid to individual differences in processing critical to decision-making aspects of comprehension.

Two studies were specifically designed to examine the correlates of individual differences in one critical aspect of decision-making -- the perception of sentence difficulty. It is clear that highly difficult material will require more detailed processing than less complicated material. Therefore, accurate perception of the underlying difficulty of the material is very important in efficient utilization of the different processing modes.

These two studies will be discussed separately below. The first deals with individual differences in perceiving syntactic structure. Since it has been shown that sentences differing in structure vary with respect to processing difficulty, it is clear that accurate perception of structure is necessary for efficient shifting of processing modes. The second study includes variables, in addition to syntactic complexity, which are thought to influence sentence difficulty (for example, word familiarity and content meaningfulness.)

B. Experiment 1: Perception of Syntactic Structure

In addition to potentially causing difficulties in making processing decisions, the inability to appropriately use grammatical structure may have other implications. In particular, Weinstein and Rabinovitch (1971) examined differences in the extent to which "good" and "poor" readers were able to take advantage of grammatical structure in learning lists of nonsense syllables. Two groups of Ss were given the task of committing to memory lists of nonsense syllables which varied in the amount of grammatical structure that was present. One kind of list was completely unstructured. A second was in the form of a pseudo-sentence; that is it contained English function words and bound morphemes such that if the nonsense syllables had been replaced by appropriate English words, grammatical sentences would have resulted. An example of this type of construction (with the bound morphemes underlined) is: When they sivoled the veg, they hanashed zaflv. A third type of list was identical to the second except for the ommission of the bound morphemes. Epstein's (1961, 1962) work demonstrated that the pseudo sentences significantly facilitated learning and the addition of bound morphemes resulted in significantly faster learning than the pseudo sentences alone. Weinstein and Rabinovitch showed that while this is true for good readers, significant differences in learning rate are not observed with readers scoring at or below the 50th percentile on the Gates Reading Test. This suggests that sensitivity to or ability to process sentence structure may be a factor in determining reading proficiency.

This finding suggests a number of possibilities, the first of which is that persons with reading difficulties, aside from vocabulary problems, may be relatively insensitive to structural information. The result of this sort of problem might be that tht person would, because of time constraints, fail to comprehend sentences, which with more time, he could have understood.

A second and more interesting possibility is that poor readers might incorrectly use structural cues and thus assign a different structure and thus an incorrect meaning to certain sentences. In effect, this is to say that the grammar of the poor reader might simply not be the same as the grammar of his society.

In order to explore these possibilities a multidimensional rank order technique was used to determine perceived proximity between all possible pairs of a set of 16 sentences. These sentences were 8 transformations derived from each of two kernels (the two kernels are paraphrases of each other). Multidimensional scaling (Carrol and Chang, 1970) was used to derive spatial representations of the sentences for both "good" and "poor" readers. These spaces were then compared to determine differences related to reading ability.

(1) Method

Subjects

Thirty-one General Psychology students served as subjects in order to fulfill the course requirement. Seventeen of these Ss were classified as "poor" readers (total reading percentages on the Cooperative English Reading Test ranging from 10% to 71%). and fourteen were classified as "good" readers (total reading percentages ranging from 77% to 99%).

Stimuli

The stimuli used in this experiment were 16 English sentences. The sentences varied from 9 to 12 words in length (depending on the transformation)and contained high familiarity words (selected from the set of "A" words in Thorndike and Loge, 1944). Of the sixteen sentences, 8 were from one kernel (K), passive (P), question (Q), negative (N), PN, PQ, NQ, and PNQ family and the other 8 were from a second K, P, N, ... PNQ family which was based on a paraphrase of the first kernel. The second kernel was derived from the first by synonym substitution. These sentences were punched and printed on IBM cards in order to facilitate the administration and analysis of the multidimensional rank order task.

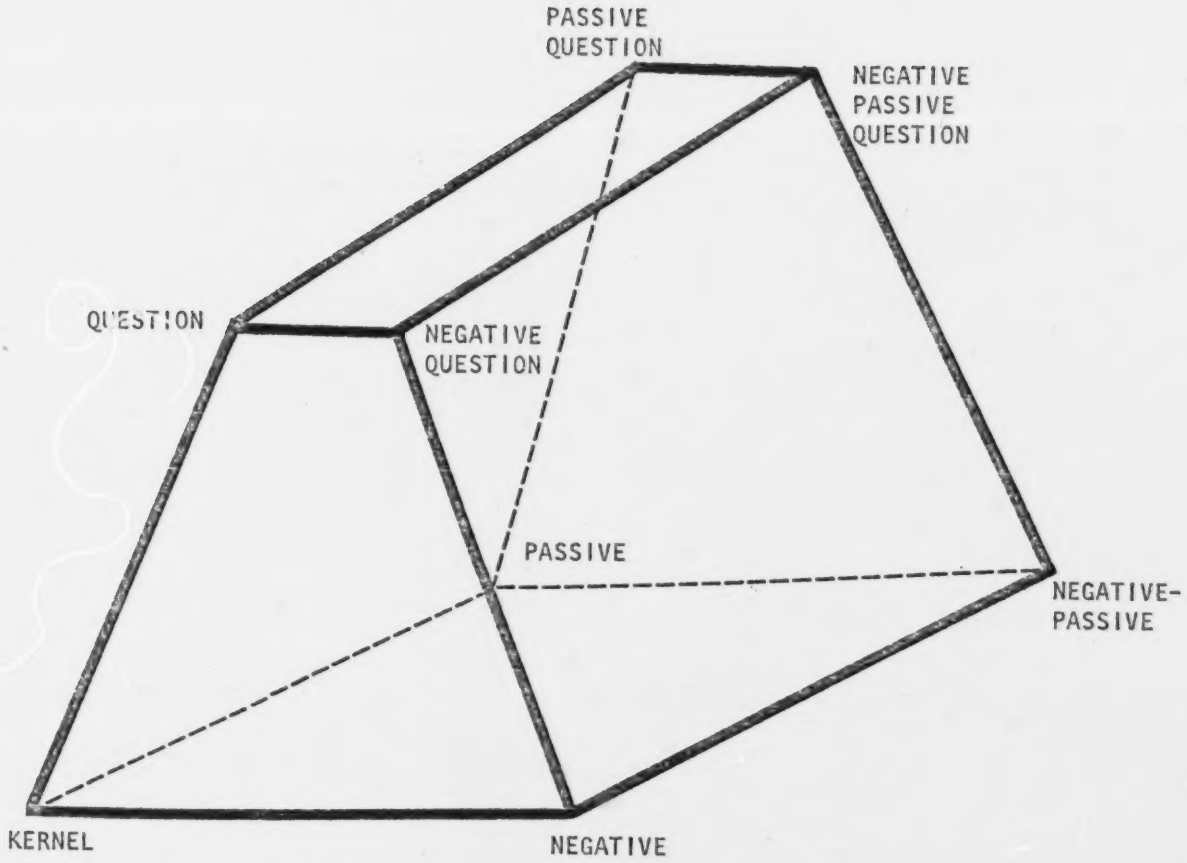
Prodedure

Each S was presented with 16 sets of 16 sentences, each sentence printed on a separate card. The first card in each set (a pink card) served as the "reference sentence" for that set. The S's task was to decide which sentence in each set was most likely to be confused with the "reference sentence". That sentence was placed immediately behind the pink card containing the "reference sentence". Next, the S selected the sentence which was the second most likely candidate for confusion and placed it behind the first sentence, and so on. Ss went through all sixteen sets of cards in this manner without time constraints (average running time was approximately 40 minutes).

(2) Results and Discussion

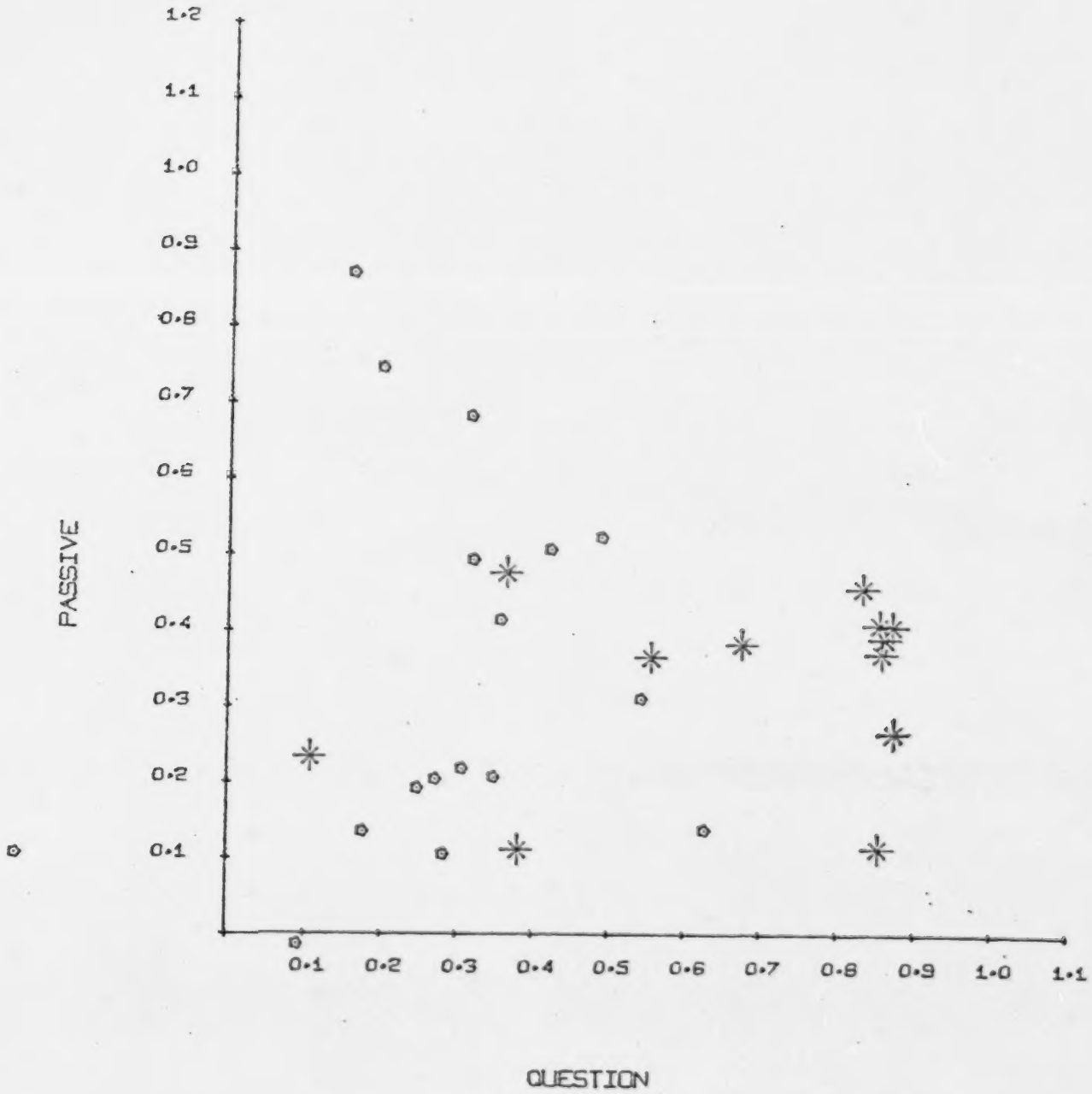
The rank-order data was analyzed by the INDSCAL multidimensional scaling program (Carrol and Chang, 1970). A four dimensional solution was judged optimal, with the following dimensions predominating: negative - non-negative, active - passive, question - non-question, and paraphrase - non-paraphrase. This solution closely replicates the findings of Clifton and Odom (1965), who found a relationship between transformations which is best characterized by Figure 4.

There are some rather striking differences between good and poor readers in performing the rank-order task. These differences are clearly shown by a series of pairwise dimension plots (Figures 5-10). In all graphs the pentagons represent poor readers, the stars represent good readers, and the values along the axes represent the weighting of the dimension in an individual's decision process. A number of observations can be made: (a) Bayesian and multiple discriminant analyses lead to highly significant separations of good and poor readers, thereby indicating substantial differences in the processing strategies employed by these groups of Ss; (b) good readers exhibit less inter-subject variability (for example, they tend to cluster more tightly) than poor readers, thus indicating strategy communalities among good readers which do not exist among poorer readers; (c) good readers weight the question dimension more in their judgments than do poor readers; and (d) good readers weight the passive dimension less than poor readers. It should be noted that these differences in performance cannot be directly attributed to differences in intelligence, since reading scores and intelligence are only negligibly correlated.



A SCHEMATIC OF THE RELATIONSHIPS
BETWEEN TRANSFORMATIONS
FOUND BY CLIFTON AND ODOM (1965)

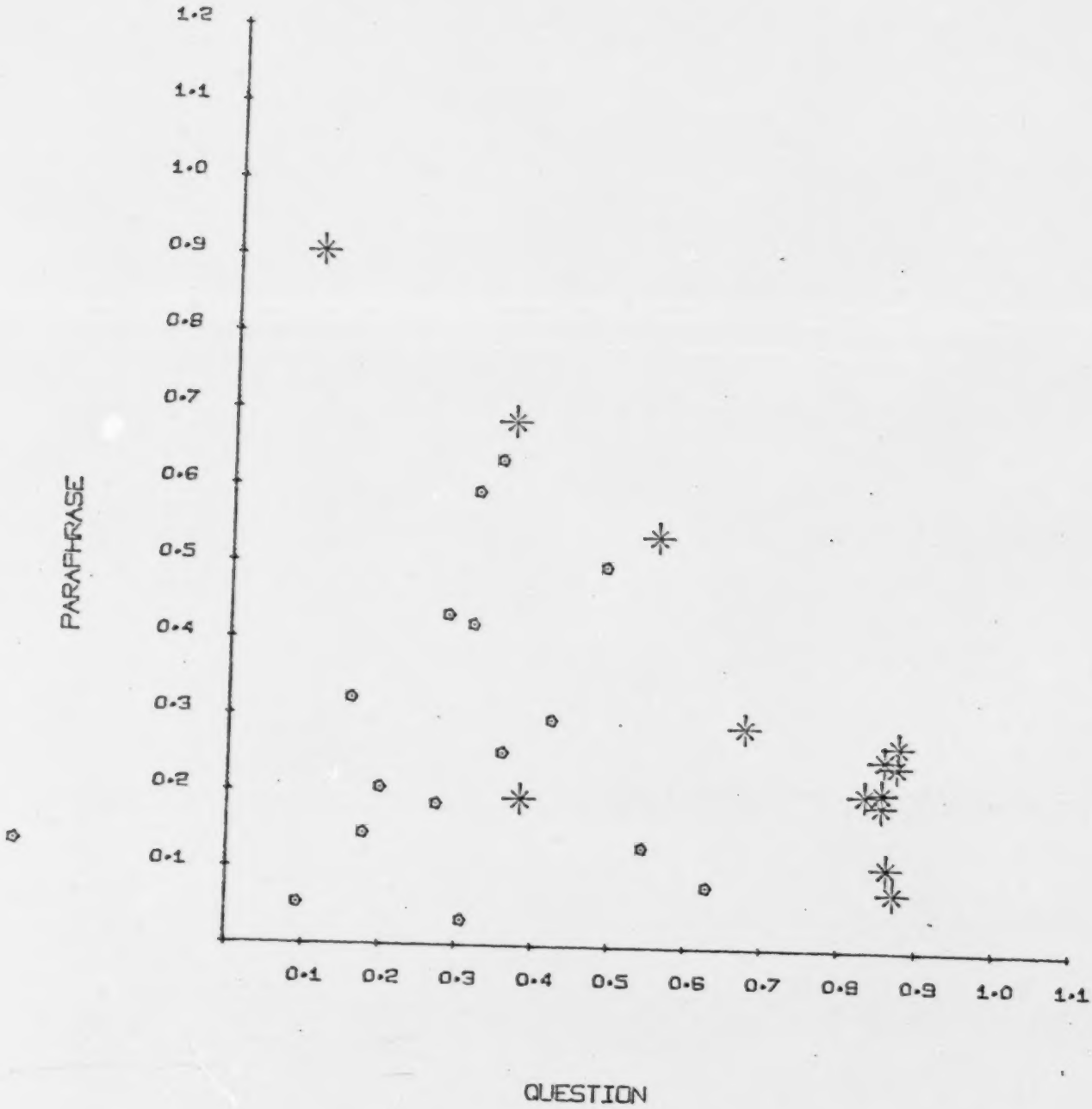
FIGURE 4



WEIGHTING OF THE PASSIVE DIMENSION
vs. WEIGHTING OF THE QUESTION
DIMENSION

✳ GOOD READERS
Ø POOR READERS

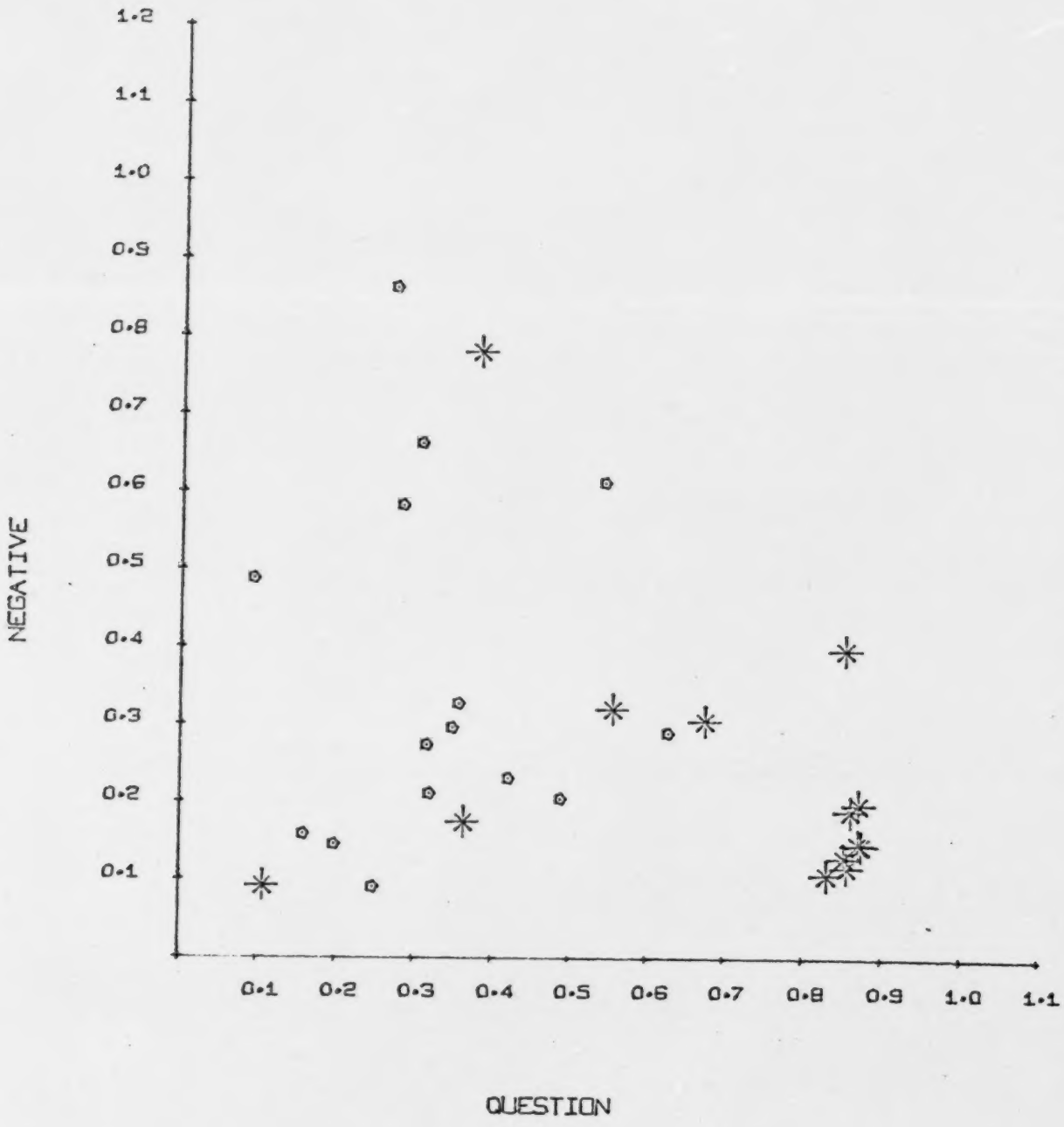
FIGURE 5



WEIGHTING OF THE PARAPHRASE DIMENSION
vs. WEIGHTING OF THE QUESTION
DIMENSION

GOOD READERS
POOR READERS

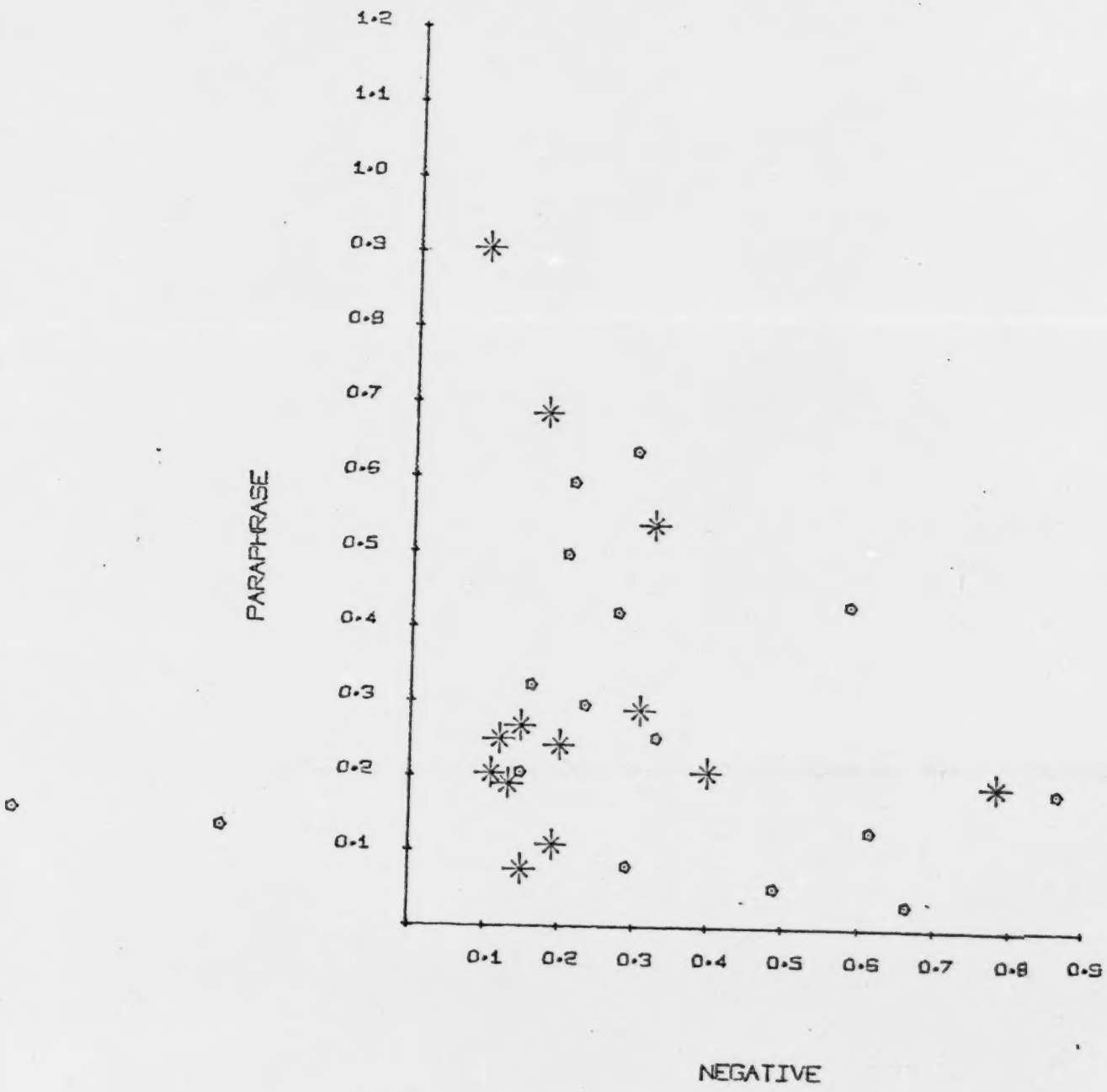
FIGURE 6



WEIGHTING OF THE NEGATIVE DIMENSION
vs. WEIGHTING OF THE QUESTION
DIMENSION

GOOD READERS
POOR READERS

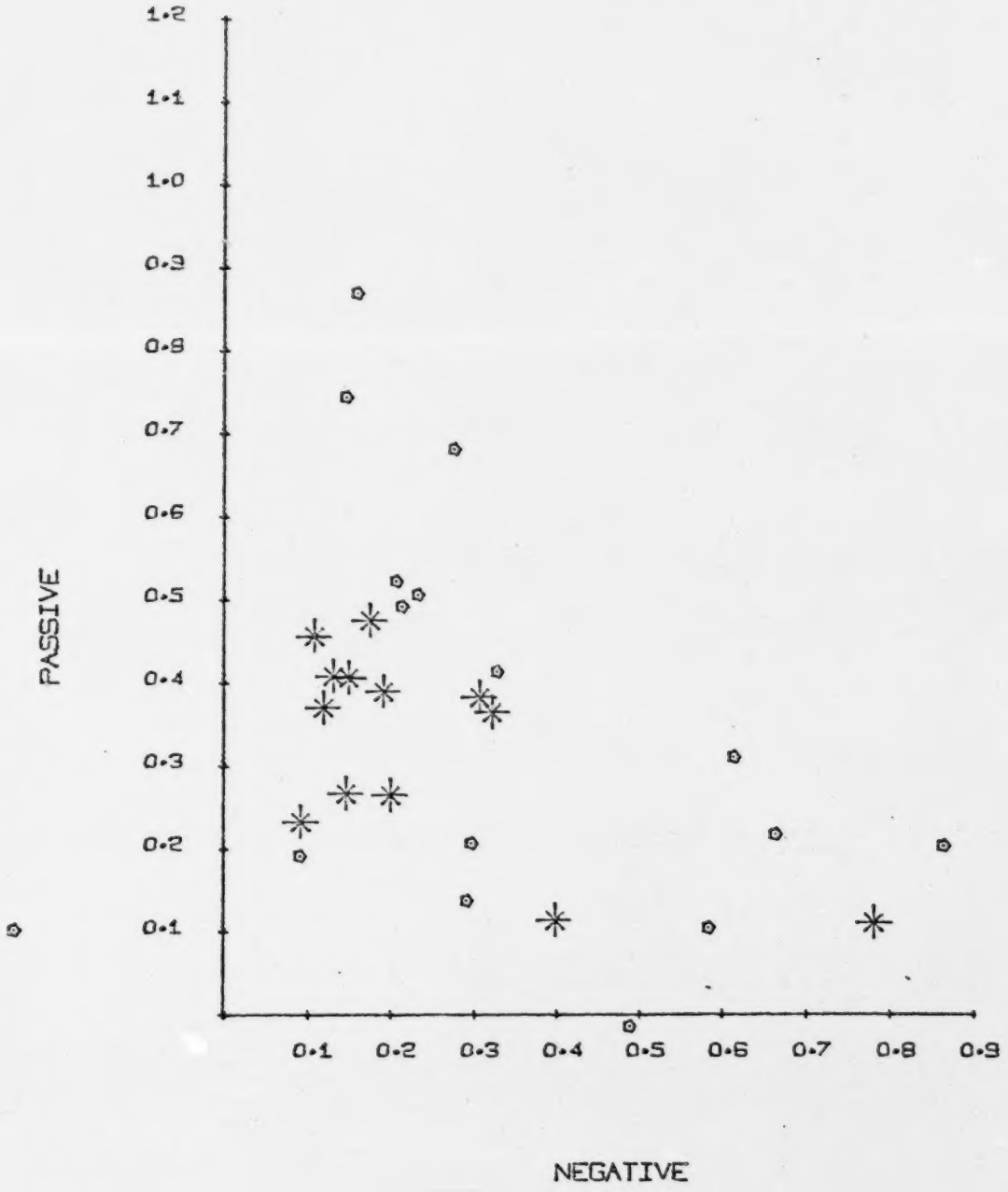
FIGURE 7



WEIGHTING OF THE PARAPHRASE DIMENSION
vs. WEIGHTING OF THE NEGATIVE
DIMENSION

✳ GOOD READERS
⬠ POOR READERS

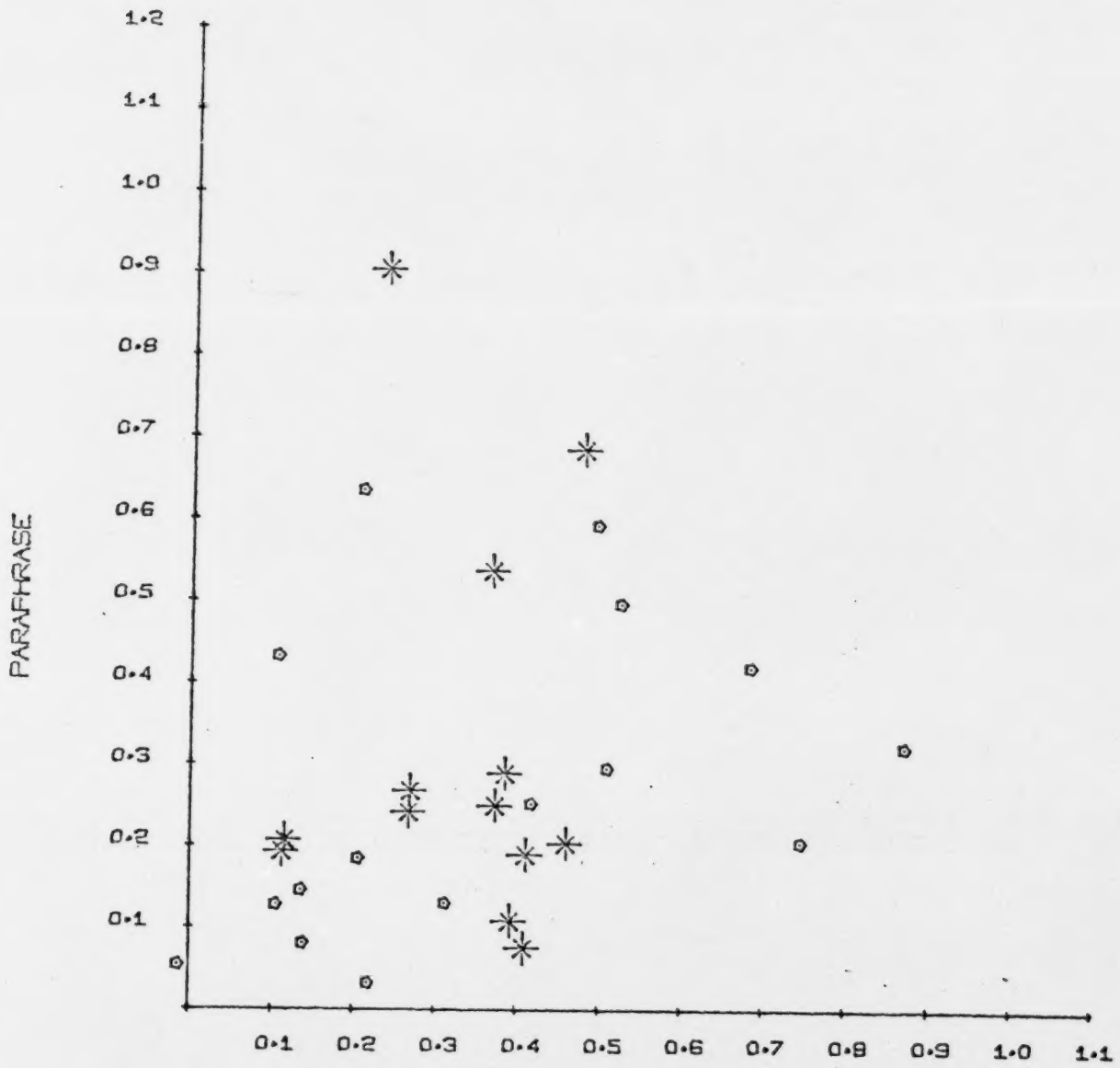
FIGURE 8



WEIGHTING OF THE PASSIVE DIMENSION
vs. WEIGHTING OF THE NEGATIVE
DIMENSION

✳ GOOD READERS
⬠ POOR READERS

FIGURE 9



WEIGHTING OF THE PARAPHRASE DIMENSION
vs. WEIGHTING OF THE PASSIVE
DIMENSION

✳ GOOD READERS
⬡ POOR READERS

FIGURE 10

(3) Conclusion

It can be concluded that good and poor readers do not perceive syntactic or grammatical structures in the same way. These results strongly support the findings of Weinstein and Rabinovitch (1971), and, therefore, imply that some reading difficulties, even at the college level, may be due to inappropriate processing of grammatical structures. In particular, these processing differences may be reflected in differences in reading flexibility.

C. Experiment 2: The Effect of Reading Ability and Exposure Time on Judgment of Complexity, Familiarity, and Comprehensibility.

The principal aim of this study was to determine if "good" and "poor" readers differ in their perception of sentence difficulty under varying time conditions. Specifically the following variables were manipulated or measured:

Sentence difficulty - Eight sets of stimulus sentences (48 in all) were produced varying in two levels of (a) syntactic complexity (active vs. passive) (Honeck, 1971), (b) word familiarity (high vs low frequency words derived from the Thorndike-Lorge word frequency list) (Pavio, et al., 1968), and (c) content meaningfulness (semantically logical vs. semantically illogical sentences) (Danks, 1969 for similar procedure). Previous work indicates that variations along these dimensions will contribute to the perceived difficulty of a sentence (Honeck, 1971; Danks, 1969; Pavio, et al., 1968). Each sentence contained from nine to eleven words with equal numbers of nouns and verbs in each. The sentences were randomly ordered for experimental presentation.

Individual differences - The thirty Ss with the lowest "total reading" percentile on the Cooperative English Reading Test were classified as "poor" readers; the thirty Ss in the highest percentiles were designated "good" readers. The arbitrary cut-off point for the lowest "good" reader score was the 75th percentile, whereas the highest "poor" reader score was at the 56th percentile.

Time considerations - In order to examine possible time contingencies which relate to decision-making ability, the stimulus sentences were presented under a fast and slow time condition. A pre-experimental test suggested that a 1.5 second exposure of each sentence would be roughly equivalent to a "scan" or "skim" processing mode; whereas a 4 second stimulus exposure would allow for more detailed processing. It was hypothesized that individual differences might be more readily detected across these two varying time conditions. That is, "poor" readers might misperceive sentence difficulty under "fast" conditions, but not under "slow" conditions, whereas "good" readers would be accurate under both conditions.

The major dependent variable in this study was subjective comprehensibility judgments. A few recent studies have indicated that comprehensibility judgments correlate quite well with objective measures of comprehensibility (Schwartz, et al., 1970). This choice of dependent measure provided us with an

opportunity to investigate another, closely related aspect of the comprehension process. Specifically, it has been pointed out by several investigators (for example, Carroll, 1971) that little is known about the subjective dimensions which underlie these judgments. Therefore, a close examination of the relation between Ss' judgments of comprehensibility and the experimental control of the aforementioned sentence parameters, might bring information to bear on this issue. Furthermore, Ss were also asked to provide judgments of syntactic complexity and content familiarity. Relationships among the dependent measures should provide further evidence as to the subjective basis of comprehensibility judgments.

(1) Method

Subjects - Sixty Introductory Psychology students at Texas Christian University served as subjects. Their participation was a course requirement.

Procedure - There were two parts to this experiment, corresponding to the previously mentioned time conditions. Sentences were mounted on slides and presented under mechanical control on a screen in front of the S. In Part I all Ss, run in groups of thirty, were given a 1.5 second exposure of each sentence, followed by a 15 second interval. During this interval they were to make judgments (on a scale from 1 - 10) of (a) syntactic complexity, (b) content familiarity, and (c) comprehensibility of the previous sentence.

Two weeks later, the same Ss were used in Part II. Part II was identical to Part I with two exceptions: (a) The exposure time for each sentence was 4 seconds instead of 1.5 second, and (b) After each stimulus presentation in Part II, Ss were asked to judge the veracity (True or False) of a paraphrase of the previous sentence. This second addition was included as a check to keep the Ss honest.

(2) Results and Discussion

At this point it is most informative to subdivide our data analyses into three categories which indicate their relationship to: the accuracy of perceiving difficulty, consistency, and the bases of the comprehensibility judgments.

Accuracy of perceiving difficulty - Two 2-way ANOVA's (corresponding to the two time conditions) have been carried out. The first factor was reading ability (good vs. poor), and the second was the within subject factor of sentence difficulty (the eight groups of sentences classified according to experi-

mental control). With comprehensibility judgments as the dependent measure, significant results were obtained on sentence difficulty, but not on reading ability. An interaction between sentence difficulty and reading ability occurred in the slower time condition (see Table 1). A plot of the mean judgment under each level of sentence difficulty, with these levels clustered and arranged in approximate order of assumed difficulty, illustrates more clearly the variation in judgments across cells. (see Figures 11 and 12).

These results suggest that both good and poor readers seem to be responding to the experimental manipulation of sentence difficulty, but that the "good" readers appear to be somewhat more sensitive to these underlying dimensions, especially "meaningfulness" (note the differences in Table 2). To provide further evidence for this notion, a series of Pearson product-moment correlations were calculated. These correlations were between Ss' comprehensibility judgments and sentence difficulty. Stimulus sentences were placed in one of four categories -- ranging from easiest to most difficult -- according to the experimental manipulations of the sentence parameters. Correlations between these "levels of difficulty" and the comprehensibility estimate of that sentence for good and poor readers under both time conditions were calculated (see Table 3). Again, all correlations are significant at $p < .01$, (N=1200). Note that good readers have somewhat higher correlations than poor readers, indicating a slightly more accurate perception of the underlying parameters.

Consistency of subject's judgments - As a means of measuring Ss judgment consistencies, Ss judgment on each sentence in Time 1 was correlated with his judgment on the same sentence in Time 2. These correlations were calculated for good and poor readers under all judgment categories (see Table 4). Of note here is the apparent greater consistency of good readers across the different time conditions.

Dimensions underlying comprehensibility judgments - A thorough inspection of Figure 12(Time 2 ANOVA graph) indicates that experimental manipulation of the meaningfulness dimension seems to have the greatest effect on comprehensibility judgments. Notice the largest deviation from a best-fitting straight line occurs when experimental control of meaningfulness is different from the familiarity and complexity dimensions (middle of graph). That is, when complexity and familiarity have "easy"values, and meaningfulness has a "hard value, the comprehensibility judgment is considerably lower than the line of best fit; in the opposite control situation, however, (complexity and familiarity are "hard", meaningfulness is "easy") comprehensibility judgments are considerably higher than the line of best fit.

TIME 1 (FAST)

SOURCE	SS	df	MS	F	EXACT P
BETWEEN Ss					
A	7.257	1	7.257	1.240	0.268
S-w/in GROUPS	362.617	62	5.848		
WITHIN Ss					
B	147.679	7	21.097	59.682	0.000*
AB	3.386	7	0.483	1.368	0.215
B x S-w/in GRPS.	153.414	434	0.353		

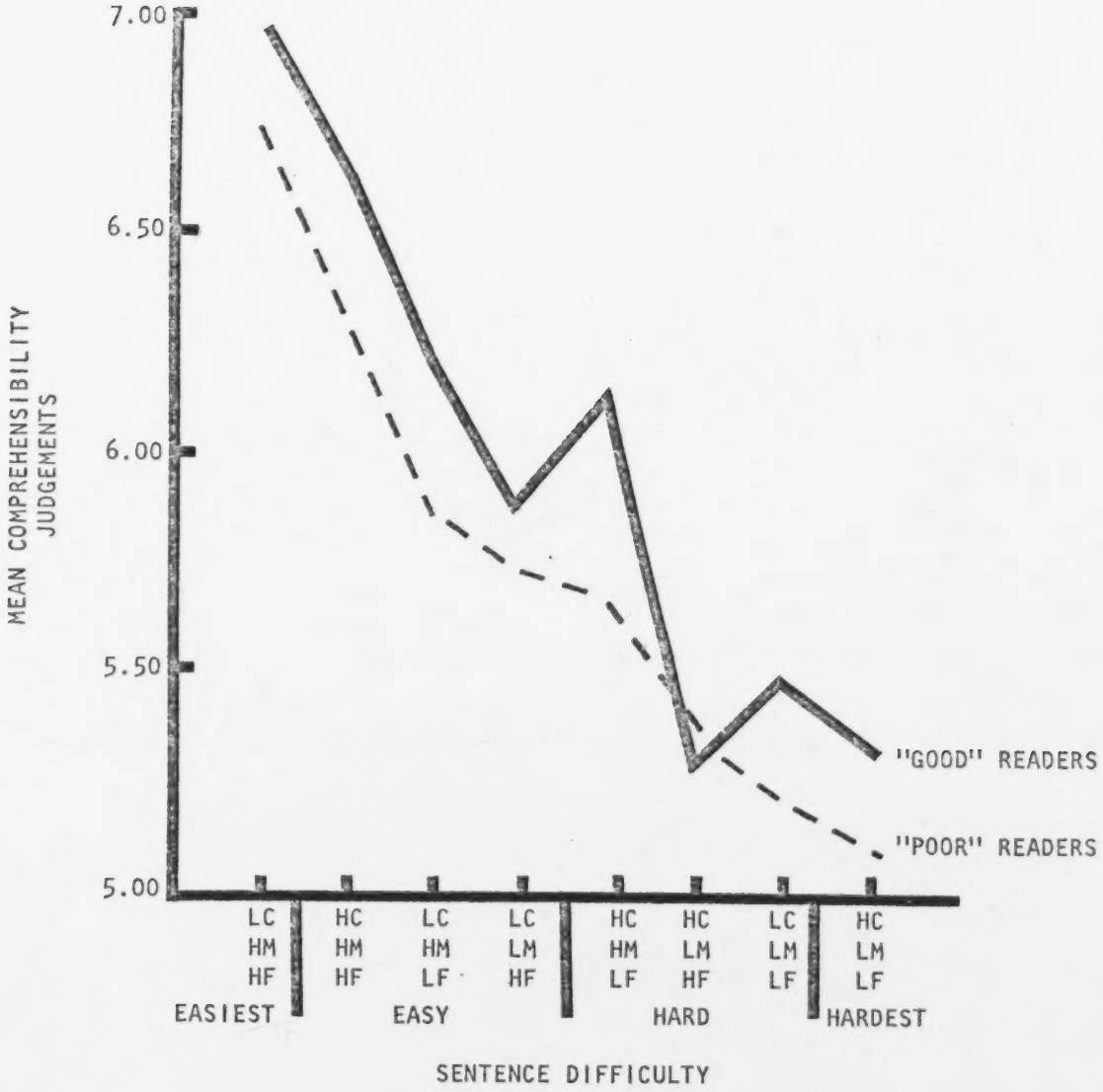
TIME 2 (SLOW)

SOURCE	SS	df	MS	F	EXACT P
BETWEEN Ss					
A	0.076	1	0.076	0.011	0.910
S-w/in GROUPS	308.951	48	6.436		
WITHIN Ss					
B	123.529	7	17.647	46.911	0.000*
AB	10.415	7	1.487	3.955	0.000*
Bx S-w/in GRPS.	126.394	336	0.376		

FACTOR A = READING ABILITY
FACTOR B = SENTENCE DIFFICULTY

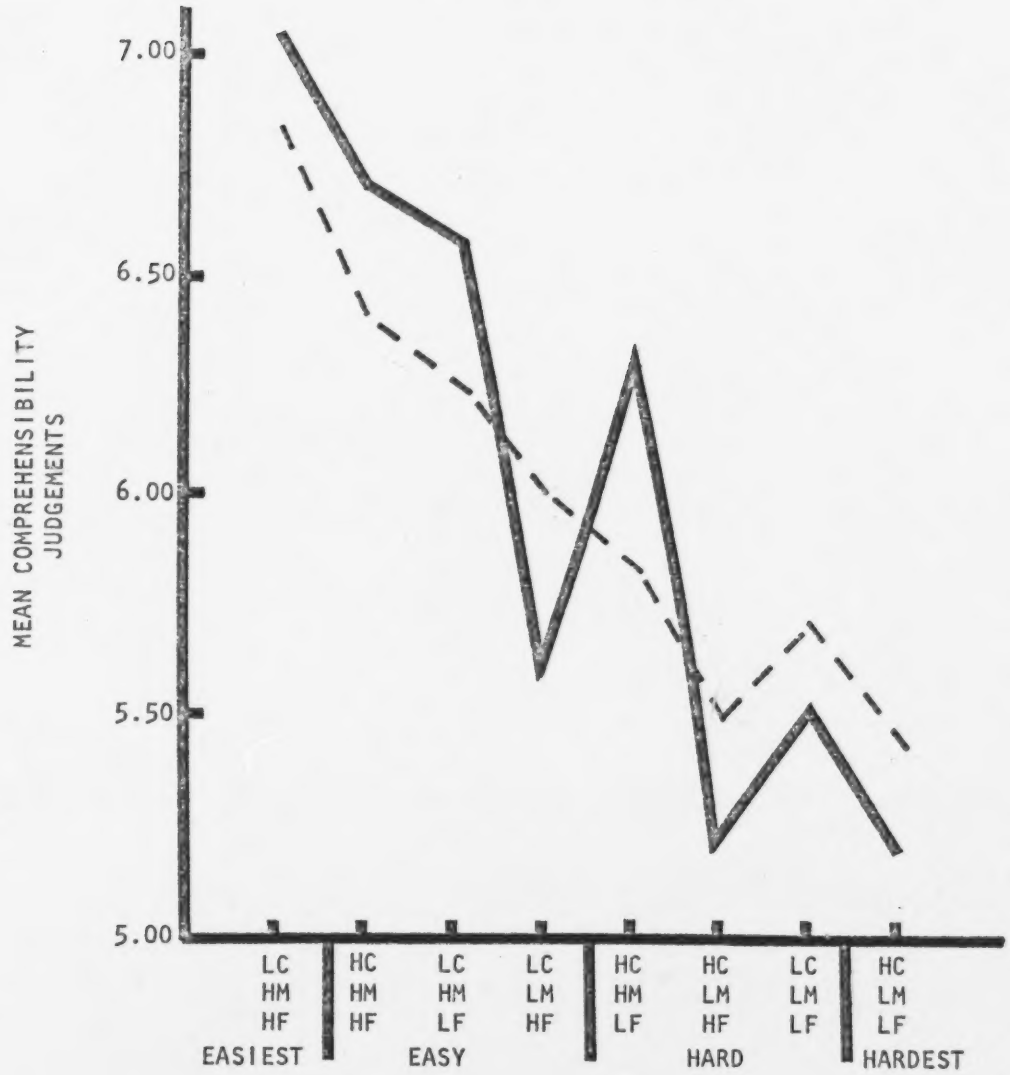
ANOVA SUMMARIES

TABLE 1



MEAN COMPREHENSIBILITY JUDGEMENTS AS A FUNCTION OF SENTENCE DIFFICULTY FOR THE TIME 1 CONDITION

FIGURE 11



MEAN COMPREHENSIBILITY JUDGEMENTS AS
A FUNCTION OF SENTENCE DIFFICULTY
FOR THE TIME 2 CONDITION

FIGURE 12

	GOOD	POOR
HIGH MEANINGFULNESS SENTENCES	6.67	6.34
LOW MEANINGFULNESS SENTENCES	5.37	5.66

MEAN COMPREHENSIBILITY JUDGMENTS

FOR VARIATIONS IN MEANINGFULNESS

TABLE 2

	GOOD	POOR
TIME 1 (FAST)	0.323	0.286
TIME 2 (SLOW)	0.347	0.273

ALL VALUES SIGNIFICANT ($p < .01$)

CORRELATIONS OF COMPREHENSIBILITY

JUDGMENTS WITH SENTENCE DIFFICULTY

TABLE 3

	GOOD	POOR
COMPLEXITY	91.3	61.9
FAMILIARITY	95.6	71.4
COMPREHENSIBILITY	86.9	66.7

PERCENTAGE OF SIGNIFICANT

CONSISTENCY CORRELATIONS

(a consistency correlation was calculated for each subject)

TABLE 4

In order to provide further evidence for this finding a series of Pearson product-moment correlations were calculated to determine the intercorrelations of the three different judgment categories (complexity, familiarity, and comprehensibility). Correlations for good and poor readers in both time conditions were obtained. Although all correlations were significant at $p < .01$, (N=1200) it is important to further note the large correlations between familiarity and comprehensibility for both good and poor readers in both time conditions (see Table 5). One highly plausible interpretation of this result is that this correlation is mediated by meaningfulness. That is, Ss perceive content familiarity and meaningfulness as being synonymous.

(3) Conclusions

In general it appears that one attribute that distinguishes "good" vs. "poor" readers is their ability to accurately and consistently perceive the difficulty of sentences. This difference in ability may manifest itself in the "flexibility" with which the readers deal with material varying in relevance and difficulty. Experiments to further delimit individual processing differences in decision-making are in the planning stage.

An additional finding from this study is the importance of meaningfulness relative to word familiarity and syntactic complexity in influencing comprehensibility judgments. This supports and expands the finding of Danks (1969). Individual differences in perception of this dimension will be explored in further studies.

TIME 1

POOR READERS

	COMPLEXITY	FAMILIARITY	COMPREHENSIBILITY
COMPLEXITY	1.00	-0.23	-0.33
FAMILIARITY		1.00	0.67
COMPREHENSIBILITY			1.00

TIME 1

GOOD READERS

	COMPLEXITY	FAMILIARITY	COMPREHENSIBILITY
COMPLEXITY	1.00	-0.10	-0.22
FAMILIARITY		1.00	0.62
COMPREHENSIBILITY			1.00

ALL VALUES SIGNIFICANT $p < .01$ (N=1200)

TIME 2

POOR READERS

	COMPLEXITY	FAMILIARITY	COMPREHENSIBILITY
COMPLEXITY	1.00	-0.27	-0.40
FAMILIARITY		1.00	0.61
COMPREHENSIBILITY			1.00

TIME 2

GOOD READERS

	COMPLEXITY	FAMILIARITY	COMPREHENSIBILITY
COMPLEXITY	1.00	-0.19	-0.27
FAMILIARITY		1.00	0.66
COMPREHENSIBILITY			1.00

JUDGMENT INTERCORRELATIONS

TABLE 5

VI. References

Bormuth, J.R. Cloze tests as a measure of comprehension ability and readability. Unpublished doctoral dissertation, Indiana University, 1962.

Bormuth, J.R. Readability: A new approach. Reading Research Quarterly, 1966, I, 79-131.

Bormuth, J.R. Comparable cloze and multiple-choice comprehension test scores. Journal of Reading, 1967, 10, 291-299.

Bormuth, J.R. Illiteracy in the suburbs. Unpublished manuscript, Department of Education, University of Chicago, 1970.

Bormuth, J.R., Manning, J., Carr, J., and Pearson, D. Children's comprehension of between - and within - sentence syntactic structures. Journal of Educational Psychology, 1970, 61, 349-357.

Bower, G., Clark, M.C., Lesgold, A.M. and Winzenz, P. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-343.

Bowman, M.E. Some relationships between flexibility and reading gains at the college level. Journal of Reading Specialist, 1966, 6 (1), 20-25.

Braam, L.S. Development and measuring flexibility in reading. The Reading Teacher, 1963, 16, 247-251.

Braam, L.S. and Berger, A. Effectiveness of four methods of increasing reading rate, comprehension, and flexibility. Journal of Reading, 1968, 11 (5), 346-352.

Brown, J.J. Evaluating student performance in listening, Education, 1955, 75, 316-321.

Carrol, J.D. and Chang, J.J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart - Young" decomposition. Psychometrika, 1970, 35 (3), 283-319.

Carroll, J.B. Learning from verbal discourse in educational media: A review of the literature. Research Bulletin RB-71-61, Educational Testing Service, Princeton, New Jersey, 1971.

Clifton, C. and Odom, P. Similarity relations among certain English sentence constructions. Psychological Monographs, 80, 5, 35pp.

Collins, A.M. and Quillian, M.R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Collins, A.M. and Quillian, M.R. Experiments on semantic memory and language comprehension. In L.W. Gregg (Ed.) Cognition in learning and memory. New York: John Wiley, 1970.

Conant, M.M. The Construction of a Diagnostic Reading Test. New York: Teachers College, Columbia University. Contributions to Education No. 861, 1942.

Crothers, E.J. Memory structure and the recall of discourse. University of Colorado, Department of Psychology. Technical Report CLIPR-4, April, 1971.

Danks, J.H. Grammaticalness and meaningfulness in the comprehension of sentences. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 687-696.

Dansereau, D.F., Fenker, R.M., and Evans, S.H. Visual pattern perception: encoding and storage of schematic versus random patterns. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 49-50.

Davis, F.B. Fundamental factors of comprehension in reading. Psychometrika, 1944, 9, 185-197.

Davis, F.B. Research in comprehension in reading. Reading Research Quarterly, 1968, 3, 499-545.

Epstein, W. The influence of syntactic structure on learning. American Journal of Psychology, 1961, 74, 80-85.

Epstein, W. A further study of the effect of syntactic structure on learning. American Journal of Psychology, 1962, 75, 121-126.

Frase, L.T. Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. Journal of Educational Psychology, 1969a, 60, 394-401.

Frase, L.T. Structural analysis of the knowledge that results from thinking about text. Journal of Educational Psychology Monograph, 1969b, 60 (6, Pt. 2).

Gough, P.B. Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-111.

Gough, P.B. The verification of sentences: The effects of delay of evidence and sentence length. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 492-496.

Greenough, D.E. and Semmel, M.I. Comprehension and imitation of sentences by Mongoloid children as a function of transformational complexity. Paper presented at the meeting of the American Educational Research Association, Los Angeles, February, 1969.

Hall, W.E. and Robinson, F.P. An analytical approach to the study of reading skills. Journal of Educational Psychology, 1945, 36, 429-442.

Honeck, R.P. A study of paraphrases. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 367-381.

Jones, S. The effect of a negative qualifier in an instruction. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 497-501.

Kintsch, W. Models for free recall and recognition. In D.A. Norman (Ed.) Models of human memory. New York: Academic Press, 1970.

Laycock, F. Significant characteristics of college students with varying flexibility in reading rate: I. Eye movements in reading prose. II. Motor and perceptual skill in "reading" material whose meaning is unimportant. Journal of Experimental Education, 1955, 23, 311-330.

Lorge, I. Predicting reading difficulty of selections for children. Elementary English Review, 1939, 16, 229-233.

MacGinitie, W.H. Contextual constraint in English prose paragraphs. Journal of Psychology, 1961, 51, 121-130.

MacGinitie, W.H. and Tretiak, R. Measures of sentence complexity as predictors of the difficulty of reading materials. Proceedings, 77th Annual Convention, American Psychological Association, 1969, 667-668.

Mandler, G. Organization and memory In K.W. Spence and J.T. Spence (Eds.) The psychology of learning and motivation. Vol. I. New York: Academic Press, 1967.

Miller, G.A. A psychological method to investigate verbal concepts. Journal of Mathematical Psychology, 1969, 6, 169-191.

Miller, G.A., Gallanter, E., and Pribram, K.H. Plans and the structure of behavior. New York: Holt, Rhinehart and Winston, 1960.

Oakan, R., Wiener, M., and Cromer, W. Identification, organization, and reading comprehension for good and poor readers. Journal of Educational Psychology, 1971, 62, 71-78.

McDonald, A.S. Research in the classroom. Journal of Reading, 1965, III (3), 187-191.

Paivio, A., Yuille, J.L., and Modigan, S.A. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology (Monograph Supplement), 1968, 76, 1-25.

Rankin, E.F. The cloze procedure - a survey of research. In L.L. Thurstone (Ed.) Yearbook, National Reading Conference. 1965, 50-93.

Rourke, D.L. Hierarchical organization in free recall. Technical Report No. 35, Perceptual Systems Laboratory, U.C.L.A., 1971.

Ruddell, R.B. The effect of the similarity of oral and written patterns of language structure on reading comprehension. Elementary English, 1965, 42, 403-410.

Segal, E.M. Hierarchical structure in free recall. Journal of Experimental Psychology, 1969, 80, 59-63.

Schwartz, D.S., Sparkman, J.P., and Deese, J. The process of understanding and judgment of comprehensibility. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 87-93.

Shipley, E.F., Smith, C.S., and Gleitman, L.R. A study in the acquisition of language: Free responses to commands. Technical Report No. 8, December, 1967, Eastern Pennsylvania Psychiatric Institute, Grant No. MH07990, National Institute of Mental Health.

Singer, H. Theoretical models of reading. The Journal of Communication, 1969, 19, 134-156.

Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.

Stoker, H.W. and Kropp, R.P. The predictive validities and factorial content of the Florida State-wide Ninth-Grade Testing Program Battery. Florida Journal of Educational Research, 1960, 105-114.

Taylor, W.L. Cloze procedure: a new tool for measuring readability. Journalism Quaterly, 1953, 30, 415-533.

Thorndike, E.L. and Lorge, I. The teacher's word book of 30,000 words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Thurstone, L.L. Note on a reanalysis of Davis' Reading Tests. Psychometrika, 1946, 11, 185-188.

Torgerson, W.S. Theory and method of scaling. New York: John Wiley, 1958.

Weinstein, R. and Rabinovitch, M.S. Sentence structure and retention in good and poor readers. Journal of Educational Psychology, 1971, 62, 25-30.